

Space Division Multiple Access with a Sum Feedback Rate Constraint

Kaibin Huang, Robert W. Heath, Jr and Jeffrey G. Andrews

Abstract

On a multi-antenna broadcast channel, simultaneous transmission to multiple users by joint beamforming and scheduling is capable of achieving high throughput, which grows double logarithmically with the number of users. The sum rate for channel state information (CSI) feedback, however, increases linearly with the number of users, reducing the effective uplink capacity. To address this problem, a novel space division multiple access (SDMA) design is proposed, where the sum feedback rate is upper-bounded by a constant. This design consists of algorithms for CSI quantization, threshold based CSI feedback, and joint beamforming and scheduling. The key feature of the proposed approach is the use of feedback thresholds to select feedback users with large channel gains and small CSI quantization errors such that the sum feedback rate constraint is satisfied. Despite this constraint, the proposed SDMA design is shown to achieve a sum capacity growth rate close to the optimal one. Moreover, the feedback overflow probability for this design is found to decrease exponentially with the difference between the allowable and the average sum feedback rates. Numerical results show that the proposed SDMA design is capable of attaining higher sum capacities than existing ones, even though the sum feedback rate is bounded.

I. INTRODUCTION

For a multi-antenna communication downlink, *space division multiple access* (SDMA) allows simultaneous transmission through the spatial separation of scheduled users. The high throughput of SDMA led to its inclusion in the IEEE 802.16e standard [1]. Compared with the optimal SDMA strategy that uses *dirty paper coding* [2], [3], SDMA with transmit beamforming has suboptimal performance but a low-complexity transmitter. Various methods for designing SDMA under beamforming constraints have been proposed recently, including zero forcing [4]–[7], a signal-to-interference-plus-noise-ratio (SINR) constraint [8], minimum mean squared error (MMSE) [9], and channel decomposition [10]. In a network, SDMA beamforming can be combined with scheduling to further improve the throughput by exploiting multi-user diversity,

Authors are with The University of Texas at Austin. Email: {khuang, rheath, jandrews }@ece.utexas.edu

This work is funded by Freescale Inc. and the National Science Foundation under grants CCF-514194.

which refers to the selection of users with good channels for transmission [11]–[17]. Typically, joint beamforming and scheduling for SDMA requires users to send back their channel state information (CSI). Therefore, given that all users share a common feedback channel, the sum feedback rate can rapidly become a bottleneck for a SDMA system with a large number users. That motivates us to address in this paper the following questions: *How to design a SDMA downlink with a bounded sum feedback rate? Does this sum feedback rate constraint significantly affect the system performance?*

A. Prior Work and Motivation

The sum feedback rate of a downlink system can be reduced by applying a feedback threshold, where users below the threshold do not send back CSI. This feedback reduction algorithm was first proposed in [18] for a downlink system with single-input-single-output (SISO) channels, where only users meeting a signal-to-noise-ratio (SNR) threshold are allowed to send back SNR information for scheduling. This algorithm is shown to reduce the sum feedback rate significantly. To further reduce the sum feedback rate, the feedback reduction algorithm in [18] is modified in [19] to have an adaptive threshold. The drawback of this modified algorithm is the feedback delay due to multiple rounds of feedback and also the additional feedback cost incurred by this process. In [20], [21], for both SISO and multiple-input-single-output (MISO) channels, combining a feedback threshold and one bit feedback per user is shown to achieve the optimal growth rate of the sum capacity with the number of users. A common problem shared by these feedback reduction algorithms is that the sum feedback rate increases linearly with the number of users, placing a burden on the uplink channel if the number of users is large.

To constrain the sum feedback rate, an approach combining a feedback threshold and contention feedback is proposed in [22] for SISO channels, where feedback users contend for the use of a common feedback channel. By extending this approach to MIMO channels, a SDMA algorithm is proposed in [23], which nevertheless has limitations for practical implementation. First, the number of simultaneous users supported by space division is limited by the number of receive antennas for each user, which is usually very small. Second, every user must inefficiently

perform zero-forcing equalization even though only a small subset of users is scheduled for transmission. These limitations motivate us to consider a more practical downlink system.

In the literature of SDMA with transmit beamforming, the sum feedback rate constraint has not been considered as most work focuses on feedback reduction for individual users. For the *opportunistic SDMA* (OSDMA) algorithm proposed in [16], the feedback of each user is reduced to a few bits by constraining the choice of a beamforming vector to a set of orthogonal vectors. The sum capacity of OSDMA can be increased by selecting orthogonal beamforming vectors from multiple sets of orthogonal vectors, which motivates the *OSDMA with beam selection* (OSDMA-BS) [15] and the *OSDMA with limited feedback* (OSDMA-LF) [24] algorithms¹. These two algorithms assign beamforming vectors at mobiles and the base station, respectively. Existing SDMA algorithms share the drawback of having a sum feedback rate that increases linearly with the number of users. This motivates us to apply a sum feedback rate constraint on SDMA.

B. Contributions and Organization

We propose an algorithm for a SDMA downlink with orthogonal beamforming and the average sum rate for CSI feedback upper-bounded by a constant, which is referred to as the sum feedback rate constraint. This constraint is enforced by using two feedback thresholds for selecting feedback users, which gives the name of the algorithm: *OSDMA with threshold feedback* (OSDMA-TF). First, a feedback threshold on users' channel power selects users with large channel gains for feedback. Second, a threshold on users' channel quantization errors prevents CSI quantization from stopping the growth of the sum capacity with the number of users [14], [24]. The key differences between OSDMA-TF and existing algorithms are summarized as follows. Contrary to the sum feedback reduction algorithms for SISO channels [18]–[21], [27] and other OSDMA algorithms with finite-rate feedback [15], [16], [24], OSDMA-TF satisfies the sum feedback rate constraint. Among downlink algorithms enforcing this constraint, OSDMA-TF has the advantage of supporting simultaneous users compared with the SISO contention feedback

¹Limited feedback refers to quantization and feedback of CSI [25], [26]

algorithm in [22] and the advantage of having simple receivers for subscribers compared with the MIMO contention feedback algorithm in [23].

The main contributions of this paper are the OSDMA-TF algorithm, the design of feedback thresholds for enforcing the sum feedback rate constraint, and the analysis of the impact of this constraint on the sum capacity. First, the OSDMA-TF sub-algorithms for *CSI quantization at users, selection of feedback users using thresholds and joint beamforming and scheduling at a base station* are proposed. Second, the feedback thresholds on users' channel power and channel quantization errors are designed such that the sum feedback rate constraint is satisfied. Third, from an upper-bound, the feedback overflow probability is found to decrease approximately exponentially with the difference between the allowable and the average sum feedback rates. Fourth, it is shown that the growth rate of the sum capacity with the number of users can be made arbitrarily close to the optimal one by having a sufficiently large sum feedback rate. Last, OSDMA-TF is compared with several existing SDMA algorithms and is found to be capable of achieving higher sum capacities despite the sum feedback rate constraint. The main conclusion of this paper is that the proposed SDMA algorithm allows a sum feedback rate constraint to be applied on a SDMA downlink without causing any appreciable negative impact.

The remainder of this paper is organized as follows. The system model is described in Section II. The algorithms used for CSI quantization, CSI feedback and joint beamforming and scheduling are presented in Section III. The feedback thresholds are derived in Section IV-A, along with an upper bound for the feedback overflow probability. In Section V, the sum capacity for OSDMA-TF is analyzed. The performance of OSDMA-TF is compared with existing SDMA algorithms using Monte Carlo simulations in Section VI, followed by concluding remarks in Section VII.

II. SYSTEM MODEL

The downlink system illustrated in Fig. 1 is described as follows. A base station with N_t antennas transmits data simultaneously to N_t scheduled users chosen from a total of U users, each with one receive antenna. The base station separates the multi-user data streams by using

beamforming, i.e. assigning a beamforming vector to each of the N_t scheduled users. The beamforming vectors $\{\mathbf{w}_n\}_{n=1}^{N_t}$ are selected from multiple sets of unitary orthogonal vectors following the beam and scheduling algorithm described in Section III-C. The received signal of the n th scheduled user is expressed as

$$y_n = \sqrt{P} \sum_{i=1}^{N_t} \mathbf{h}_n^\dagger \mathbf{w}_i x_i + \nu_n, \quad n = 1, \dots, N_t, \quad (1)$$

where we use the following notation

N_t number of transmit antennas and also number of scheduled users;

\mathbf{h}_n ($N_t \times 1$ vector) downlink channel;

P transmit power;

\mathbf{w}_n ($N_t \times 1$ vector) beamforming vector with $\|\mathbf{w}_n\|^2 = 1$;

x_n transmitted symbol with $|x_n| = 1$;

y_n received symbol; and

ν_n AWGN sample with $\nu_n \in \mathcal{CN}(0, 1)$.

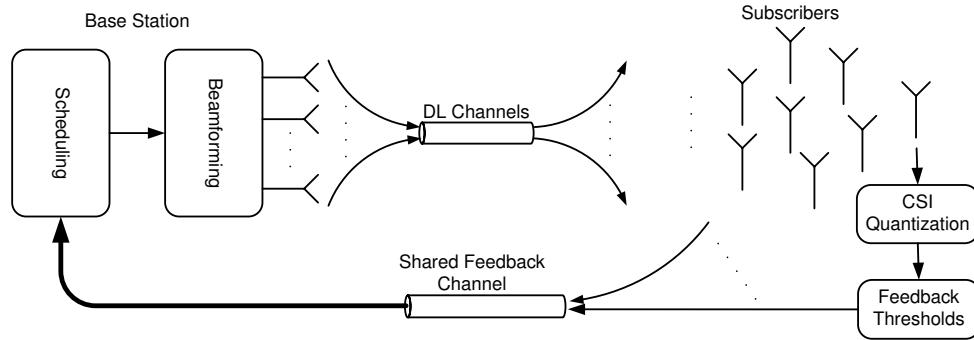


Fig. 1. SDMA Downlink system with feedback thresholds

We assume that each user quantizes his/her CSI and sends it back following a feedback algorithm to be discussed in Section III-B. Furthermore, all users share a common feedback channel. Therefore, it is necessary to constrain the average sum feedback rate. Let B denote the number of bits sent back by each feedback user and K the number of feedback users. It follows that the instantaneous sum feedback rate is BK . Since B is a constant and K a random

variable, the constraint of the average sum feedback rate can be written as

$$(\text{Sum Feedback Rate Constraint}) \quad BE[K] \leq R, \quad (2)$$

where R is the sum feedback rate constraint.

To simplify the analysis of the proposed algorithms in Section III, we make the following assumption about the multi-user channels:

AS 1: *The downlink channel \mathbf{h}_u is an i.i.d. vector whose coefficients are $\mathcal{CN}(0, 1)$.*

Given this assumption, which is commonly made in the literature of SDMA and multi-user diversity [12], [13], [15], [16], [28], the channel direction vector of each user follows a uniform distribution, which greatly simplifies the design of feedback thresholds in Section IV-A and capacity analysis in Section IV-A.

III. ALGORITHMS

OSDMA-TF is comprised of (i) CSI quantization at the subscribers, (ii) selection of feedback users using feedback thresholds, and (iii) joint beamforming and scheduling at the base station. The algorithms for performing these functions are discussed in Section III-A to Section III-C, respectively. Furthermore, OSDMA-TF is compared with existing SDMA algorithms in Section III-D.

A. CSI Quantization

Without loss of generality, the discussion in this section is focused on the u th user and the same algorithm for CSI quantization is used by other users. For simplicity, we assume:

AS 2: *The u th user has perfect receive CSI \mathbf{h}_u .*

This assumption allows us to neglect channel estimation error at the u th mobile. For convenience, the CSI, \mathbf{h}_u , is decomposed into two components: the *gain* and the *shape*, which are quantized separately. Hence, $\mathbf{h}_u = g_u \mathbf{s}_u$ where $g_u = \|\mathbf{h}_u\|$ is the gain and $\mathbf{s}_u = \mathbf{h}_u / \|\mathbf{h}_u\|$ is the shape. The channel shape \mathbf{s}_u is quantized and sent back to the base station for choosing beamforming vectors. The channel gain g_u is used for computing SINR, which is also quantized and sent back

as a channel quality indicator. Due to the ease of quantizing SINR that is a scalar, we make the following assumption:

AS 3: *The SINR is perfectly known to the base station through feedback.*

The same assumption is made in [15], [16]. This assumption allows us to focus our discussion on quantization of the channel shape \mathbf{s}_u .

Quantization of the channel shape \mathbf{s}_u is the process of matching it to a member of a set of pre-determined vectors, called *a codebook*. Different from [24], [28] where code vectors are randomly generated, we propose a structured codebook constructed as follows. The codebook, denoted as \mathcal{F} , is comprised of M sub-codebooks: $\mathcal{F} = \cup_{m=1}^M \mathcal{F}_m$, each of which is comprised of N_t orthogonal vectors. The sub-codebooks $\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_M$ are independently and randomly generated for example using the method in [29]. Each sub-codebook provides a potential set of orthogonal beamforming vectors for downlink transmission. Given a codebook \mathcal{F} thus generated, the quantized channel shape, denoted as $\hat{\mathbf{s}}_u$, is the member of \mathcal{F} that forms the smallest angle with the channel shape \mathbf{s}_u [30]. Mathematically,

$$\hat{\mathbf{s}}_u = \mathcal{Q}(\mathbf{s}_u) = \arg \max_{\mathbf{f} \in \mathcal{F}} |\mathbf{f}^\dagger \mathbf{s}_u|, \quad (3)$$

where the function \mathcal{Q} represents the CSI quantization process. We define the *quantization error* as

$$(\text{Quantization Error}) \quad \delta_u = \sin^2(\angle(\hat{\mathbf{s}}_u, \mathbf{s}_u)). \quad (4)$$

It is clear that the quantization error is zero if $\hat{\mathbf{s}}_u = \mathbf{s}_u$.

B. Feedback Algorithm

To satisfy the sum feedback rate constraint (2), we propose a threshold-based feedback algorithm, which allows only users with good channels (high SINRs) to send back their CSI to the base station. The SINR for the u th user is a function of the channel power $\rho_u = \|\mathbf{h}_u\|^2$ and the quantization error δ_u in (4) (see also [24]):

$$\text{SINR}_u = \frac{1 + P\rho_u}{1 + P\rho_u\delta_u} - 1. \quad (5)$$

Therefore, the feedback algorithm employs two feedback thresholds for feedback user selection: the channel power threshold, denoted as γ , and the quantization error threshold, denoted as ϵ . It follows that the u th user meets the feedback criteria if $\rho_u \geq \gamma$ and $\delta_u \leq \epsilon$. The thresholds γ and ϵ are designed in Section IV-A such that the sum feedback rate constraint in (2) is satisfied.

Given that the u th user meets the feedback thresholds, the quantized channel shape $\hat{\mathbf{s}}_u$ is sent back to the base station through a finite-rate feedback channel [25], [30]². Since the quantization codebook \mathcal{F} can be known *a priori* to both the base station and mobiles, only the index of $\hat{\mathbf{s}}_u$ needs to be sent back. Therefore, the number of feedback bits per user is $\log_2 N$ since $|\mathcal{F}| = N$.

C. Joint Beamforming and Scheduling

Among feedback users, the base station schedules a subset of users for downlink transmission using the criterion of maximizing sum capacity and under the constraint of orthogonal beamforming. To facilitate the description of the procedure for joint beamforming and scheduling, we group feedback users according to their quantized channel shapes by defining the following index sets:

$$\mathcal{I}_{m,n} = \{1 \leq u \leq U \mid \rho_u \geq \gamma, \delta_u \leq \epsilon, \mathcal{Q}(\mathbf{s}_u) = \mathbf{f}_{m,n}\}, \quad 1 \leq m \leq M, 1 \leq n \leq N_t, \quad (6)$$

where $\mathcal{Q}(\cdot)$ is the quantization function in (3) and $\mathbf{f}_{m,n} \in \mathcal{F}$ is the n th member in the m th sub-codebook $\mathcal{F}_m \subset \mathcal{F}$. The base-station adopts a two-step procedure for joint beamforming and scheduling. First, it selects the user with maximum SINR from each index set defined (6). Second, from these selected users, the base station schedules up to N_t users for downlink transmission under the constraint of orthogonal beamforming. Following this procedure, the resultant sum capacity can be written as

$$\mathcal{C} = E \left[\max_{m=1, \dots, M} \sum_{n=1}^{N_t} \log_2 \left(1 + \max_{u \in \mathcal{I}_{m,n}} \text{SINR}_u \right) \right], \quad (7)$$

where the two “max” operators correspond to the two steps in the procedure for joint beamforming and scheduling. In the event that an index set $\mathcal{I}_{m,n}$ is empty, we set $\max_{u \in \mathcal{I}_{m,n}} \text{SINR}_u = 0$ in (7).

²The feedback of SINR is ignored due to AS 3.

D. Comparison with Existing Algorithms

We summarize in Table I the key differences between OSDMA-TF and existing SDMA algorithms with all-user feedback, including OSDMA-LF, OSDMA-BS and OSDMA. Performance comparisons by Monte Carlo simulation are provided in Section VI.
[h]

TABLE I

COMPARISON OF OSDMA-TF, OSDMA-LF, OSDMA-S AND OSDMA

	OSDMA-TF	OSDMA-LF	OSDMA-BS	OSDMA
# of Feedback Users	NN_t	U	U	U
Feedback/User (bits) ^a	$2B + \log_2 N$	$2B + \log_2 N$	$B + I \log_2 N_t$	$B + \log_2 N_t$
Sum Capacity (bits/s/Hz) ^b	largest (7.5)	largest (7.5)	moderate (6.4)	smallest (6.2)
Beamforming & Scheduling	centralized	centralized	distributed	N/A

^a Assume B bits are required for quantizing a channel gain and the quantization error of the channel shape.

^b Sum capacity is computed for $U = 20$, $N_t = 4$ and SNR = 10dB. Following [15], the sum capacity is reduced by the feedback overhead factor $\alpha = 5\%$ for each round of CSI feedback.

^c Refer to possibility that different users select a same beamforming vector.

IV. FEEDBACK DESIGN

In Section IV-A, the feedback thresholds for OSDMA-TF (cf. Section III-B) are designed as functions of the number of users under the sum feedback constraint in (2). Even if this constraint is satisfied, it is likely that the instantaneous sum feedback rate exceeds the maximum allowable feedback rate of the feedback channel and hence causes an overflow. In Section IV-B, an upper-bound for the feedback overflow probability is derived, which is useful for designing the maximum feedback rate for the feedback channel.

A. Feedback Thresholds

The feedback probability of each user is derived as a function of the feedback thresholds. Subsequently, since the sum feedback rate is proportional to this probability, we can thus derive the feedback thresholds for a given sum feedback rate.

The feedback probability of a user is defined as the probability that the user's channel power and quantization error meet the respective thresholds. We focus on the feedback probability of a single user since the channels of different users are i.i.d. given AS 1 and hence the feedback probabilities of different users are identical. For simplicity, we omit the user index, hence the subscript u , for all notation in this section. Given AS 1, the channel power $\rho = \|\mathbf{h}\|^2$ and the channel direction $\mathbf{s} = \mathbf{h}/\|\mathbf{h}\|$ are independent. It follows that the two events, namely the channel power and quantization error thresholds are met, are independent. Therefore, we can derive their probabilities separately. First, the probability that the channel power ρ of a user meets the power threshold is obtained as

$$P_\gamma = \Pr\{\rho \geq \gamma\} = \int_\gamma^\infty f_\rho(\rho) d\rho, \quad (8)$$

where $f_\rho(\rho)$ is the chi-squared PDF function given as

$$f_\rho = \frac{\rho^{L-1} e^{-\rho}}{(L-1)!}. \quad (9)$$

Second, the probability for meeting the quantization error threshold, denotes as P_ϵ , is obtained. To this end, we define a set for each member of the codebook \mathcal{F} as

$$\mathcal{V}_n = \{\|\mathbf{v}\| = 1 \mid 1 - |\mathbf{v}^H \mathbf{f}_n|^2 \leq \epsilon\}, \quad n = 1, 2, \dots, N. \quad (10)$$

Intuitively, the set \mathcal{V}_n can be viewed as a “cone” with the radius ϵ and the axis \mathbf{f}_n . The probability, P_ϵ , can be defined in terms of these sets as

$$P_\epsilon = \Pr\{\mathbf{s} \in \cup_n \mathcal{V}_n\}. \quad (11)$$

By applying the union bound and using the symmetry of different users' channels,

$$P_\epsilon \leq \sum_{n=1}^N \Pr\{\mathbf{s} \in \mathcal{V}_n\} = N \Pr\{\mathbf{s} \in \mathcal{V}_n\}. \quad (12)$$

By denoting $1 - |\mathbf{s}^H \mathbf{f}_n|^2$ as δ and from (10),

$$P_\epsilon \leq N \Pr\{\delta \leq \epsilon\} = N \epsilon^{L-1}, \quad (13)$$

where we use the following result from [28],

$$\Pr\{\delta \leq \epsilon\} = \epsilon^{L-1}. \quad (14)$$

By combining (8) and (13), the feedback probability for each user is given in the following lemma.

Lemma 1: The feedback probability for each user is given as

$$P_v = P_\gamma P_\epsilon \leq N \epsilon^{L-1} \int_\gamma^\infty f_\rho(\rho) d\rho. \quad (15)$$

The sum feedback rate, denoted as R , can be expressed as $R = E[K]B$ where $E[K]$ denotes the average number of feedback users and B the number of bits sent back by each of them. Furthermore, $E[K]$ can be written as

$$E[K] = UP_v, \quad (16)$$

with P_v given in (15). With B fixed, the sum feedback rate is proportional to $E[K]$. We derive a set of feedback thresholds such that $E[K]$ is limited by an upper-bound, which is independent of the number of users U . By choosing a proper value for the upper-bound, we can thus satisfy any given constraint on the sum feedback rate R . These results are shown as the following theorem.

Theorem 1: Consider the following channel power and quantization error thresholds

$$\gamma = \log U - \lambda \log \log U, \quad \lambda > 0, \quad (17)$$

$$\epsilon = [U^{1-\varphi} (\log U)^{\varphi \lambda}]^{-1/(L-1)}, \quad (18)$$

where

$$\varphi = -\gamma \ln \left(\frac{1}{L!} \int_\gamma^\infty \rho^{L-1} e^{-\rho} d\rho \right). \quad (19)$$

Given these thresholds, the average number of feedback users $E[K]$ is upper-bounded as

$$E[K] \leq NN_t, \quad (20)$$

where N is the cardinality of the CSI quantization codebook \mathcal{F} .

Proof: The theorem follows by substitution of the feedback thresholds in (17) and (18) into (15) and then (16). \square

A few remarks are in order:

- Given the feedback thresholds in Theorem 1, the sum feedback rate is bounded as

$$R \leq BN N_t, \quad (21)$$

where $B = B_s + \log_2 N$ is the number of feedback bits per user with B_s is the number of bits for quantizing the SINR feedback³ (cf. Section III).

- The power and quantization thresholds in (17) and (18) are chosen jointly to ensure the capacity of each scheduled user grows with the number of users U at an optimal rate, namely $\log_2 \log_2 U$ [16]. Detailed analysis is given in Section V.
- The parameter λ in (17) and (18) affects the signal-to-interference ratio (SIR) of feedback users. Its optimal value for maximizing sum capacity can be chosen via numerical methods since analytical methods seem difficult.
- CSI quantization error causes interference between simultaneous users and can potentially prevent the sum capacity from increasing with the number of users as observed in [28], [31]. This motivates the design of the quantization error threshold in (18). This threshold ensures that the quantization error of each feedback user converges to zero with the number of users U . This result is proved shortly.

To prove that the quantization errors of feedback users diminishes with the number of users U , we require the Alzer's bounds for the gamma function rewritten as the following lemma [32]

Lemma 2 (Alzer's Inequality): *The incomplete Gamma function is bounded as*

$$[1 - e^{-\beta\gamma}]^{N_t} < \int_0^\gamma f_\rho(\rho) d\rho < [1 - e^{-\gamma}]^{N_t}, \quad (22)$$

where $\beta = (N_t!)^{-1/N_t}$ and $f_\rho(\rho)$ is the chi-squared PDF in (9).

Using this lemma, we obtain the result as shown in the following corollary of Theorem 1.

Corollary 1: *The quantization error threshold ϵ in (18) ensures the quantization error of a feedback user, δ , converges to zero with the number of users U*

$$\lim_{U \rightarrow \infty} \delta \leq \lim_{U \rightarrow \infty} \epsilon = 0. \quad (23)$$

Proof: See Appendix A. □

³Usually, $B_s \ll \log_2 N$ since SINR is a scalar while the channel direction is a vector.

Last, we provide the following proposition, which shows that the upper-bound on the average number of feedback users $E[K]$ is tight if the number of users U is large. We define the minimum distance between any two members of the codebook \mathcal{F} as

$$\Delta\delta_{\min} = \min_{1 \leq a, b \leq N} [1 - |\mathbf{f}_a^H \mathbf{f}_b|^2], \quad (24)$$

where $\mathbf{f}_a \in \mathcal{F}$ and $\mathbf{f}_b \in \mathcal{F}$.

Proposition 1: *For any codebook \mathcal{F} with $\Delta\delta_{\min} > 0$, there exists an integer U_0 such that $\forall U \geq U_0$, the average number of feedback users $E[K]$ is given as*

$$E[K] = NN_t. \quad (25)$$

Proof: See Appendix B. □

For illustration of Theorem 1 and Proposition 1, the numbers of feedback users $E[K]$ averaged over different channel realizations and randomly generated codebooks are plotted against different numbers of users U in Fig. 2. First, $E[K]$ is observed to be upper-bounded by NN_t , which agrees with Theorem 1. Second, $E[K]$ converges to NN_t with the number of users U , which verifies Proposition 1.

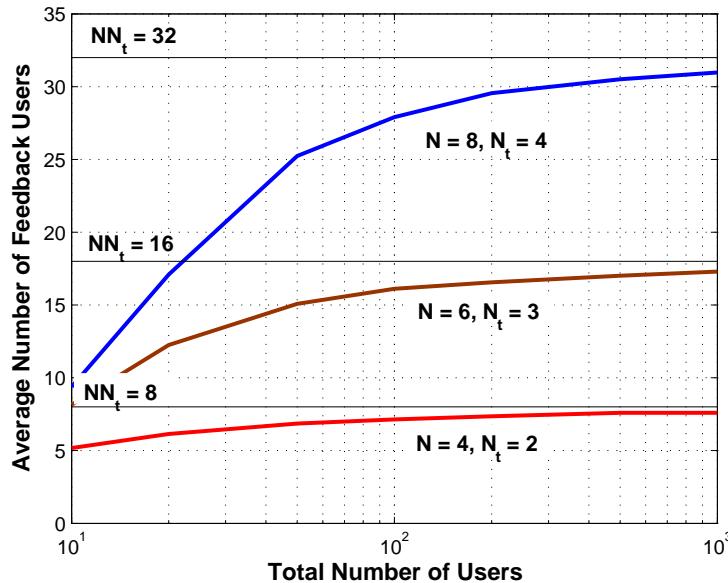


Fig. 2. Average numbers of feedback users for OSDMA-TF

B. Overflow Probability for Feedback Channel

The overflow probability is defined as the probability that the instantaneous sum feedback rate exceeds the average sum feedback rate. A small overflow probability reduces the average waiting time of feedback users and improves the system stability [33]. In this section, we show that an arbitrarily small overflow probability can be maintained by making the maximum allowable feedback rate sufficiently large relative to the average sum feedback rate.

The multiuser feedback channel satisfying the sum feedback rate constraint can be implemented *asynchronously* or *synchronously*. For the first case, the contention feedback method as in [22], [23] can be applied, which allows feedback users to compete for uplink transmission. For the second case, multiuser feedback is coordinated by a base station following a multiple-access scheme, such as *orthogonal frequency division multiple access* (OFDMA), *time division multiple access* (TDMA), or *code division multiple access* (CDMA) [34]. For both a synchronous and an asynchronous feedback channel, a small feedback overflow probability is desirable for the reasons stated earlier.

For simplicity and due to their equivalence, we measure the *instantaneous*, *average* and *maximum allowable* sum feedback rate using the instantaneous, average and maximum allowable numbers of feedback users, denoted as K , $E[K]$ and K_{\max} , respectively. The overflow probability can be upper-bounded using the Chernoff bound [35] as follows. For each user, we define a Bernoulli random variable T_u indicating whether the user meets the feedback thresholds

$$T_u = 1\{\delta_u \leq \epsilon \text{ and } \rho_n \geq \gamma\}, \quad u = 1, 2, \dots, U. \quad (26)$$

The instantaneous number of feedback user, K , can be expressed as the sum of these Bernoulli random variables, hence $K = \sum_{u=1}^U T_u$. Using the Chernoff bound for the summation of i.i.d. Bernoulli random variables derived in [35], we can obtain an upper-bound for the overflow probability.

Proposition 2: *The overflow probability of the feedback channel is upper-bounded as*

$$\Pr(K \geq K_{\max}) \leq \exp \left[-K_{\max} \log \left(\frac{K_{\max}}{E[K]} \right) - (U - K_{\max}) \log \left(\frac{U - K_{\max}}{U - E[K]} \right) \right], \quad (27)$$

where K_{\max} is the maximum number of feedback users supported by the feedback channel and $E[K] \leq NN_t$.

The upper-bound obtained above is useful for determining the maximum data rate the feedback channel should support such that a constraint on the overflow probability is satisfied since the feedback data rate is proportional to the number of feedback users.

In Fig. 3, the upper bound in (27) is compared with the actual overflow probability obtained by Monte Carlo simulation. It can be observed that both the overflow probability and its upper-bound decreases at the same slope and approximately exponentially with the difference $(K_{\max} - E[K])$.

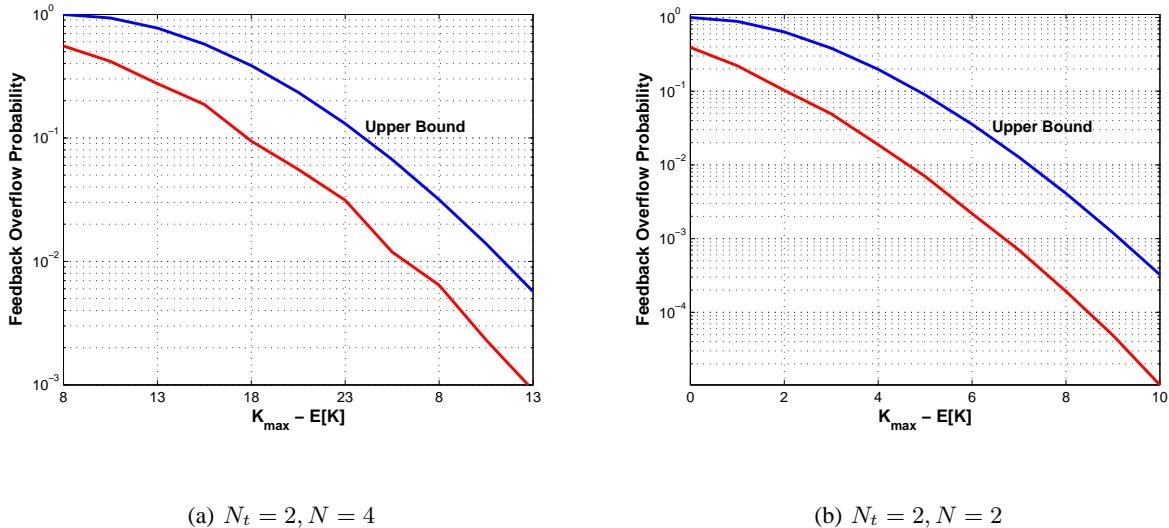


Fig. 3. Feedback channel overflow probability for OSDMA-TF

V. ANALYSIS OF SUM CAPACITY

In this section, for a large number of users ($U \rightarrow \infty$), we show that the sum capacity of OSDMA-TF can grow at a rate close to the optimal one, namely $N_t \log_2 \log_2 N$, if the sum feedback rate is sufficiently large.

Before proving the main result, several useful lemmas are presented. As we know, the high sum capacity of SDMA is due to its ability of supporting up to N_t simultaneous users. The first

lemma concerns the probability for the impossibility of scheduling N_t users. This probability is name *probability of scheduled user shortage* and denoted as P_β . Using the index sets defined in (6), we can express P_β as

$$P_\beta = \Pr \left\{ \max_{1 \leq m \leq M} \sum_{n=1}^{N_t} \mathbb{1}\{\mathcal{I}_{m,n} \neq \emptyset\} < N_t \right\}. \quad (28)$$

It can be upper-bounded as shown in Lemma 3.

Lemma 3: *The probability of scheduled user shortage is upper-bounded as*

$$P_\beta < (N_t e^{-N_t})^M. \quad (29)$$

Proof: See Appendix C. □

As to be shown later, the probability P_β characterizes the decrease of the asymptotic growth rate of the sum capacity caused by scheduled user shortage or equivalently the sum feedback rate constraint.

From (5), the multi-user interference encountered by a scheduled user with channel power ρ and channel quantization error δ is $P\rho\delta$. Lemma 4 shows that the average of the multi-user interference converges to zero with the number of user U .

Lemma 4: *Let ρ and δ denote the channel power and quantization error of a feedback user.*

We have

$$\lim_{U \rightarrow \infty} E[\rho\delta \mid \rho \geq \gamma, \delta \leq \epsilon] = 0, \quad \text{if } \lambda \geq N_t - 1, \quad (30)$$

where λ is the parameter of the power threshold in (17).

Proof: See Appendix D. □

The main result of this section is the following theorem.

Theorem 2: *For a large number of users ($U \rightarrow \infty$), the sum capacity of OSDMA-TF grows with the number of transmit antennas N_t linearly and with the number of users double logarithmically*

$$1 \geq \lim_{U \rightarrow \infty} \frac{\mathcal{C}}{N_t \log_2 \log_2 U} > 1 - (N_t e^{-N_t})^M, \quad \text{if } \lambda \geq N_t - 1, \quad (31)$$

where λ is the parameter of the power threshold in (17).

Proof: See Appendix E. □

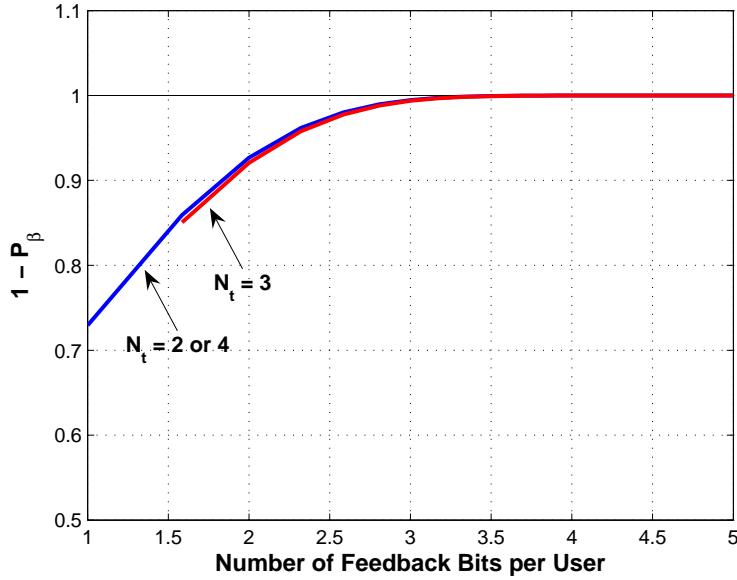


Fig. 4. Convergence of the lower bound in (31), $1 - P_\beta$, to one with the number of feedback bits per user $\log_2 N$.

The above theorem shows the effect of a sum feedback rate constraint is to decrease the growth rate of the sum capacity with respect to that for feedback from all users, namely $N_t \log_2 \log_2 U$ [16], [24]. Nevertheless, such difference in growth rate can be made arbitrarily small by increasing the sum feedback rate, or equivalently the number of feedback bits per feedback user, as stated in the following corollary.

Corollary 2: *By increasing the number of feedback bits per feedback user ($\log_2 N$), the sum capacity of OSDMA-TF can grow at the optimal rate:*

$$\lim_{N \rightarrow \infty} \lim_{U \rightarrow \infty} \frac{\mathcal{C}}{N_t \log_2 \log_2 U} = 1, \quad \text{if } \lambda \geq N_t - 1. \quad (32)$$

Proof: The result follows from (31) and $N = MN_t$. \square

We can observe from Fig. 4 that the asymptotic growth rate for the sum capacity for OSDMA-TF converges to the optimal value, hence $1 - P_\beta \rightarrow 1$ from (31), very rapidly as the number of feedback bits per user ($\log_2 N$) increases.

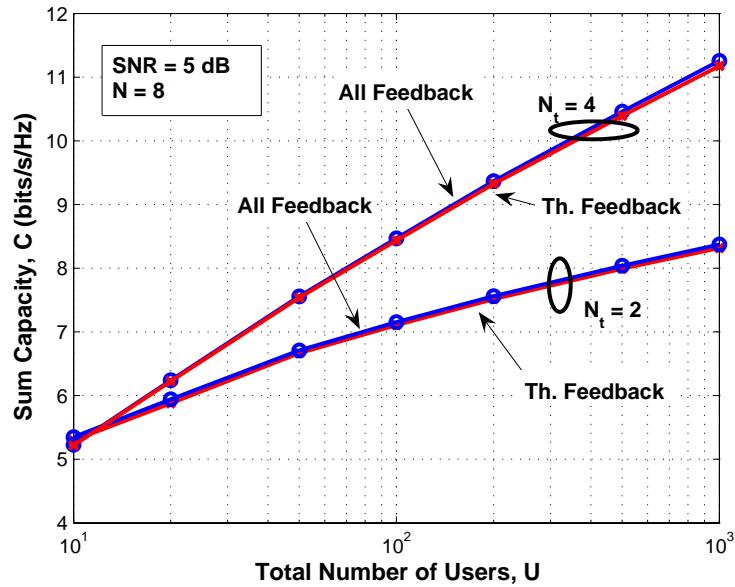
VI. PERFORMANCE COMPARISON

In this section, we compare the sum capacity and the sum feedback rate of OSDMA-TF with the case of all-user feedback, which is equivalent to OSDMA-TF with trivial feedback thresholds $\gamma = 0$ and $\epsilon = 1$. Similar comparisons are also conducted between OSDMA-TF and existing algorithms including OSDMA-LF [24], OSDMA-BS [15] and OSDMA [16].

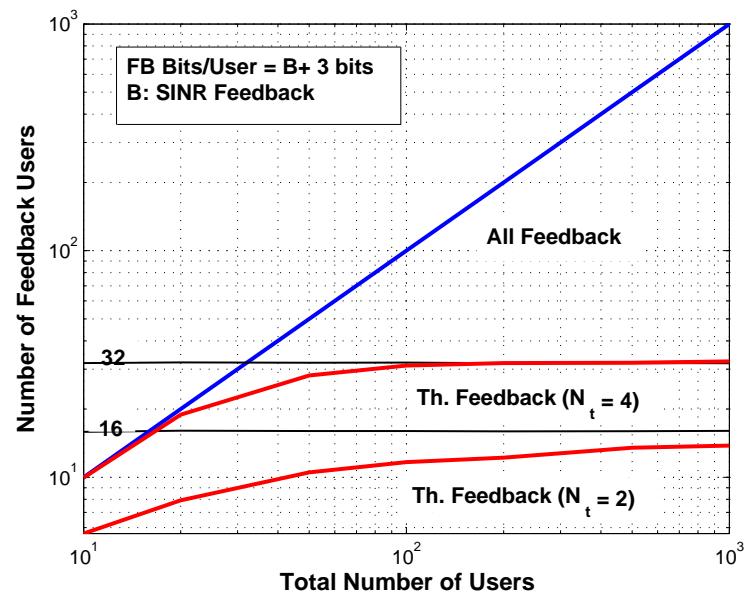
The sum capacities of OSDMA-TF and the corresponding case of all-user feedback are plotted against the number of users U in Fig. 5(a) for the cases of $N_t = \{2, 4\}$ transmit antennas. For these two cases, the parameter λ for the power threshold γ in (17) is assigned the values of 1 and 1.5 respectively, which are found numerically to be sum capacity maximizing. Each user quantizes his/her channel shape using a codebook of size $N = 8$ and hence each feedback user sends back $\log_2 N = 3$ bits. It can be observed from Fig. 5(a) that the sum feedback rate constraint for SDMA-TF incurs negligible loss in sum capacity with respect to the case of all-user feedback. Next, the number of feedback users for OSDMA-TF and all-user feedback are compared in Fig. 5(b). Note that the sum feedback rate R is proportional to the number of feedback users $E[K]$: $R = 3E[K]$ bits. It can be observed that the number of feedback users for OSDMA-TF is upper bounded by 32 for $N_t = 4$ and 16 for $N_t = 2$ since OSDMA-TF is designed for satisfying a sum feedback constraint (cf. Section IV-A). In summary, OSDMA-TF achieves almost identical sum capacity as the case of all user feedback but with a dramatic reduction on sum feedback rate for a large number of users U .

In Fig. 6, the sum capacity and sum feedback rate of OSDMA-TF is compared with those of OSDMA-LF, OSDMA-BS and OSDMA for different numbers of users U , with $N_t = 2$ and an SNR of 5 dB. The number of feedback bits per feedback user differs for the algorithms in comparison since they use different sizes for quantization codebooks or different feedback algorithms. For OSDMA-TF, two codebook sizes $N = 8$ and $N = 24$ are considered, corresponding to 3 and 4.6 feedback bits for each feedback user, respectively⁴. For OSDMA-LF, the codebook size N increases with U as: $N = 5\lceil(\log_2 U)^{N_t-1}\rceil$ to avoid saturation of sum

⁴The average number of feedback users \bar{K} is a function of N (cf. Section IV-A).



(a) Sum Capacity



(b) Sum Feedback Rate

Fig. 5. Sum capacities of threshold feedback (OSDMA-TF) and all-user feedback

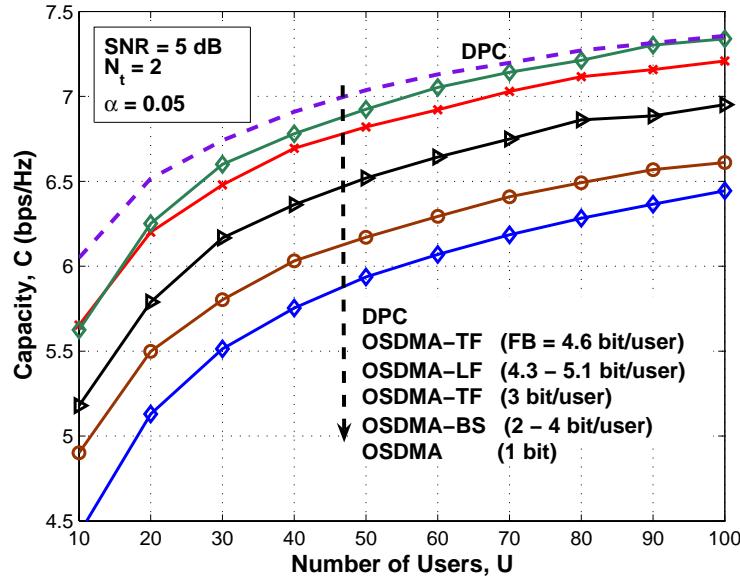
capacity due to limited feedback [24]. The codebook sizes for OSDMA-BS and OSDMA are both $N = 2$. Different from other algorithms, the CSI feedback for OSDMA-BS is performed iteratively, where each iteration penalizes the sum capacity by a factor of $0 \leq \alpha \leq 1$ [15]. Let I denote the number of feedback iterations. Therefore, for OSDMA-BS, the total feedback for each user is I bits and the sum capacity with feedback penalty is given as $\mathcal{C}_p = (1 - I\alpha)\mathcal{C}$. For fair comparison, we also apply this feedback penalty to the other algorithms⁵.

From Fig. 6(a), we can observe that OSDMA-TF ($N = 24$) yields the highest sum capacity and OSDMA-TF ($N = 8$) is outperformed only by OSDMA-LF. Moreover, the sum capacity of OSDMA-TF converges to DPC rapidly. From Fig. 6(b), the sum feedback rates for OSDMA-TF is observed to grow much more gradually with the number of users U than those for other algorithms. Asymptotically, the sum feedback rates for OSDMA-TF saturate due to sum feedback rate constraint (cf. Theorem 1) while those for other algorithms continue to increase with U . Several other observations can be made from Fig. 6(b). First, for $U \geq 25$, OSDMA-TF with $N = 8$ has the smallest sum feedback rate among all algorithms but it outperforms OSDMA-BS and OSDMA in terms of sum capacity. Second, for $U \geq 75$, OSDMA-TF with $N = 24$ yields the highest sum capacity among all algorithms and also requires the smallest sum feedback rate. Third, the sum feedback rate for OSDMA-LF is the highest among all algorithms. Fourth, the drop of sum feedback rate for OSDMA-BS at $U = 90$ is due to the decrease of the optimal number of feedback iterations, which is obtained in [15].

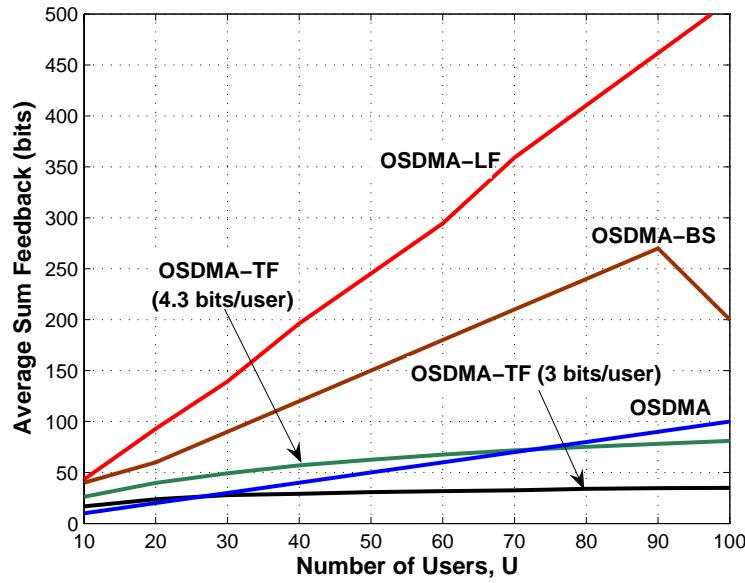
VII. CONCLUSION

This paper proposes a SDMA downlink algorithm with a sum feedback rate constraint, which is applied by using feedback thresholds on users' channel power and channel quantization errors. We derived the expressions for these thresholds and the upper bound for the corresponding feedback overflow probability. Furthermore, we obtained the asymptotic growth rate of the sum capacity with the number of users. We showed that it can be made arbitrarily close to the optimal value by increasing the sum feedback rate. From numerical results, we found that limiting the

⁵ $I = 1$ for OSDMA-TF, OSDMA-LF and OSDMA



(a) Sum Capacity



(b) Sum Feedback Rate

Fig. 6. Sum capacities and sum feedback rates of OSDMA-TF, OSDMA-LF, OSDMA-BS and OSDMA.

sum feedback rate incurs negligible loss on sum capacity. Moreover, we demonstrated that the proposed SDMA algorithm is capable of outperforming existing algorithms despite having a much smaller sum feedback rate.

APPENDIX

A. Proof of Corollary 1

From Lemma 2 and (19), we have

$$(1 - e^{-\beta\gamma})^{N_t} < 1 - N_t e^{-\varphi\gamma} < (1 - e^{-\gamma})^{N_t}. \quad (33)$$

From the definition in (17), we observe that γ monotonically increases with U when U is large. Therefore, from (17) and (33), there exists an integer U_0 such that $\forall U \geq U_0$,

$$1 - N_t e^{-\beta\gamma} < 1 - N_t e^{-\varphi\gamma} < 1 - N_t e^{-\gamma}. \quad (34)$$

It follows that $\beta < \varphi < 1$. Combining this with (18), the result in Corollary 1 follows.

B. Proof of Proposition 1

We observe from (16) and (15) that

$$E[K] = UP_\gamma P_\epsilon \leq UNP_\gamma \Pr\{\mathbf{s} \in \mathcal{V}_n\}, \quad (35)$$

where $1 \leq n \leq N$ is arbitrary. We will prove the existence of an integer U_0 such that the equality in the above equation holds $\forall U \geq U_0$. Following (18), such an integer exists such that

$$\epsilon \leq \frac{\Delta\delta_{\min}}{2}, \quad \forall U \geq U_0, \quad (36)$$

where $\Delta\delta_{\min} > 0$ is the minimum distance for the codebook \mathcal{F} as defined in (24). Assume that there exist two overlapping sets \mathcal{V}_a and \mathcal{V}_b , $\mathcal{V}_a \cap \mathcal{V}_b \neq \emptyset$. Let $\mathbf{s} \in \mathcal{V}_a \cap \mathcal{V}_b$. From the triangular inequality and the definition in (24),

$$(1 - |\mathbf{s}^H \mathbf{f}_a|^2) + (1 - |\mathbf{s}^H \mathbf{f}_b|^2) \geq \Delta\delta_{\min}. \quad (37)$$

On the other hand, from the definition in (10),

$$(1 - |\mathbf{s}^H \mathbf{f}_a|^2) + (1 - |\mathbf{s}^H \mathbf{f}_b|^2) \leq 2\epsilon. \quad (38)$$

Nevertheless, (36) leads to the contradiction between (37) and (38). Thereby, we prove that given (36),

$$\mathcal{V}_a \cap \mathcal{V}_b = \emptyset, \quad \forall U \geq U_0 \text{ and } 1 \leq a, b \leq N. \quad (39)$$

It follows that

$$\Pr\{\mathbf{s} \in \cup_n \mathcal{V}_n\} = N \Pr\{\mathbf{s} \in \mathcal{V}_1\}. \quad (40)$$

Therefore,

$$K = UNP_\gamma \Pr\{\mathbf{s} \in \mathcal{V}_1\}, \quad \forall U \geq U_0. \quad (41)$$

Substitution of (17) and (18) into the above equation completes the proof.

C. Proof of Lemma 3

From (28),

$$\begin{aligned} P_\beta &= \Pr \left\{ \bigcap_{m=1}^M \left\{ \sum_{n=1}^{N_t} 1\{\mathcal{I}_{m,n} \neq \emptyset\} < N_t \right\} \right\}, \\ &\stackrel{(a)}{=} \prod_{m=1}^M \Pr \left\{ \sum_{n=1}^{N_t} 1\{\mathcal{I}_{m,n} \neq \emptyset\} < N_t \right\}, \\ &= \prod_{m=1}^M \Pr \left\{ \bigcup_{n=1}^{N_t} \{\mathcal{I}_{m,n} = \emptyset\} \right\}, \\ &\stackrel{(b)}{\leq} \prod_{m=1}^M (N_t \Pr\{\mathcal{I}_{m,n} = \emptyset\}), \\ &\stackrel{(c)}{=} \prod_{m=1}^M [N_t (1 - P_\gamma \Pr\{\delta \leq \epsilon\})^U], \\ &\stackrel{(d)}{=} [N_t (1 - N_t/U)^U]^M, \\ &\leq [N_t e^{-N_t}]^M. \end{aligned} \quad (42)$$

The equality (a) results from the independence of the M events in (42) due to the independent generations of the M sub-codebook in the codebook \mathcal{F} . The inequality (b) is obtained by applying the union bound as well as using the equal probabilities of the events $\{\mathcal{I}_{m,n} = \emptyset\}$ for $m = 1, 2, \dots, M$. The equality (c) follows from the definition of the set $\mathcal{I}_{m,n}$ in (6). The equality (d) is obtained from (8), (14), (17) and (18).

D. Proof of Lemma 4

Given AS 1, ρ and δ are independent, hence $E[\rho\delta \mid \rho \geq \gamma, \delta \leq \epsilon] = E[\rho \mid \rho \geq \gamma]E[\delta \mid \delta \leq \epsilon]$.

By definition,

$$E[\rho \mid \rho \geq \gamma] = \frac{\int_{\gamma}^{\infty} \rho \cdot \rho^{N_t-1} e^{-\rho} d\rho}{\int_{\gamma}^{\infty} \rho^{N_t-1} e^{-\rho} d\rho} = \frac{\Gamma(N_t + 1, \gamma)}{\Gamma(N_t, \gamma)},$$

where $\Gamma(\cdot, \cdot)$ denotes the incomplete Gamma function [32]. By expanding $\Gamma(\cdot, \cdot)$, we obtain an upper-bound for $E[\rho \mid \rho \geq \gamma]$ as

$$\begin{aligned} E[\rho \mid \rho \geq \gamma] &= \frac{N_t! e^{-\gamma} \sum_{i=0}^{N_t} \gamma^i / i!}{(N_t - 1)! e^{-\gamma} \sum_{i=0}^{N_t-1} \gamma^i / i!}, \\ &= N_t \left(1 + \frac{\gamma^{N_t} / N_t!}{\sum_{i=0}^{N_t-1} \gamma^i / i!} \right), \\ &< N_t \left(1 + \frac{\gamma^{N_t} / N_t!}{\gamma^{N_t-1} / (N_t - 1)!} \right), \\ &= N_t + \gamma. \end{aligned} \quad (43)$$

Next, we obtain the expression of $E[\delta \mid \delta \leq \epsilon]$ as:

$$E[\delta \mid \delta \leq \epsilon] = \int_0^{\epsilon} \delta f_{\delta}(\delta \mid \delta < \epsilon) d\delta, \quad (44)$$

$$= \epsilon^{-(N_t-1)} \int_0^{\epsilon} \delta^L d\delta, \quad (45)$$

$$= \frac{N_t - 1}{N_t} \epsilon. \quad (46)$$

From (43) and (46),

$$0 \leq E[\rho\delta] < (N_t - 1)\epsilon + \frac{N_t - 1}{N_t} \gamma \epsilon. \quad (47)$$

From (18) and $\varphi < 1$,

$$\lim_{U \rightarrow \infty} \epsilon = 0. \quad (48)$$

Moreover, from (18) and (17),

$$\gamma \epsilon = U^{\frac{\varphi-1}{N_t-1}} (\log_2 U)^{1-\frac{\varphi\lambda}{N_t-1}} \left(1 - \lambda \frac{\log_2 \log_2 U}{\log_2 U} \right). \quad (49)$$

If $\lambda \geq N_t - 1$, it follows that

$$\lim_{U \rightarrow \infty} E[\gamma \epsilon] = 0, \quad \text{if } \lambda \geq N_t - 1. \quad (50)$$

Combining (47), (48) and (50) completes the proof.

E. Proof of Theorem Sum Capacity

The lower and upper bounds of the asymptotic sum capacity given in (31) are proved in Section E.1 and Section E.2, respectively.

1) *Lower Bound for Asymptotic Sum Capacity:*

$$m^* = \arg \max_{1 \leq m \leq M} \sum_{n=1}^{N_t} 1\{\mathcal{I}_{m,n} \neq \emptyset\}. \quad (51)$$

It follows that

$$K_{\max}^{DL} = \sum_{n=1}^{N_t} 1\{\mathcal{I}_{m^*,n} \neq \emptyset\}. \quad (52)$$

From (7):

$$\mathcal{C} \geq E \left[\sum_{n=1}^{N_t} \log_2 \left(1 + \max_{u \in \mathcal{I}_{m^*,n}} \text{SINR}_u \right) \right], \quad (53)$$

$$\geq E \left[\sum_{n=1}^{N_t} \log_2 \left(1 + \frac{1}{|\mathcal{I}_{m^*,n}|} \sum_{u \in \mathcal{I}_{m^*,n}} \text{SINR}_u \right) \right]. \quad (54)$$

By using the definition in (29) and expanding the expectation in (54),

$$\mathcal{C} \geq E \left[\sum_{n=1}^{N_t} \log_2 \left(1 + \frac{1}{|\mathcal{I}_{m^*,n}|} \sum_{u \in \mathcal{I}_{m^*,n}} \text{SINR}_u \right) \mid K_{\max}^{DL} = N_t \right] (1 - P_\beta), \quad (55)$$

$$+ E \left[\sum_{n=1}^{N_t} \log_2 \left(1 + \frac{1}{|\mathcal{I}_{m^*,n}|} \sum_{u \in \mathcal{I}_{m^*,n}} \text{SINR}_u \right) \mid K_{\max}^{DL} < N_t \right] P_\beta, \quad (56)$$

$$\geq E \left[\sum_{n=1}^{N_t} \log_2 \left(1 + \frac{1}{|\mathcal{I}_{m^*,n}|} \sum_{u \in \mathcal{I}_{m^*,n}} \text{SINR}_u \right) \mid K_{\max}^{DL} = N_t \right] (1 - P_\beta). \quad (57)$$

Since the function $\log_2(\cdot)$ is convex, it follows from (57) that

$$\mathcal{C} \geq \sum_{n=1}^{N_t} \log_2 \left(1 + E \left[\frac{1}{|\mathcal{I}_{m^*,n}|} \sum_{u \in \mathcal{I}_{m^*,n}} \text{SINR}_u \mid K_{\max}^{DL} = N_t \right] \right) (1 - P_\beta). \quad (58)$$

By substituting (5) into (58),

$$\mathcal{C} \geq \sum_{n=1}^{N_t} \log_2 \left(1 + E \left[\frac{1}{|\mathcal{I}_{m^*,n}|} \sum_{u \in \mathcal{I}_{m^*,n}} \frac{1 + P\rho_u}{1 + P\rho_u\delta_u} \mid K_{\max}^{DL} = N_t \right] \right) (1 - P_\beta). \quad (59)$$

To simplify the above expression, we use the fact that the sequences $\{\rho_u\}_{u=1}^U$ and $\{\delta_u\}_{u=1}^U$ are i.i.d., respectively. Let SINR_0 denote a random variable having the same distribution as each member of the sequence SINR_u . Then (58) can be re-written as

$$\mathcal{C} = (1 - P_\beta)N_t \log_2 \left(E \left[\frac{1 + P\rho_1}{1 + P\rho_1\delta_1} \mid \rho_1 \geq \gamma, \delta_1 \leq \epsilon \right] \right) \quad (60)$$

$$\geq (1 - P_\beta)N_t \log_2 \left((1 + P\gamma)E \left[\frac{1}{1 + P\rho_1\delta_1} \mid \rho_1 \geq \gamma, \delta_1 \leq \epsilon \right] \right). \quad (61)$$

Since the function $\frac{1}{x}$ for $x > 0$ is convex, it follows from the above inequality that

$$\mathcal{C} \geq (1 - P_\beta)N_t \{ \log_2(\gamma) + \log_2(P) - \log_2(1 + PE[\rho_1\delta_1 \mid \rho_1 \geq \gamma, \delta_1 \leq \epsilon]) \}. \quad (62)$$

By substituting (17),

$$\mathcal{C} \geq (1 - P_\beta)N_t \left[\log_2 \log_2 U + \log_2 \left(1 - \frac{\lambda \log_2 \log_2 U}{\log_2 U} \right) + \log_2(P) - \right. \quad (63)$$

$$\left. \log_2(1 + PE[\rho\delta]) \right]. \quad (64)$$

Therefore,

$$\lim_{U \rightarrow \infty} \frac{\mathcal{C}_U}{N_t \log_2 \log_2 U} \geq (1 - P_\beta)[1 + \Pi_1 + \Pi_2 + \Pi_3] \quad (65)$$

where

$$\Pi_1 = \lim_{U \rightarrow \infty} \frac{1}{\log_2 \log_2 U} \log_2 \left(1 - \frac{\lambda \log_2 \log_2 U}{\log_2 U} \right), \quad (66)$$

$$\Pi_2 = \lim_{U \rightarrow \infty} \frac{\log_2(P)}{\log_2 \log_2 U}, \quad (67)$$

$$\Pi_3 = \lim_{U \rightarrow \infty} \frac{1}{\log_2 \log_2 U} \log_2(1 + PE[\rho\delta]). \quad (68)$$

The values of Π_1 and Π_2 can be observed to be zeros. The value of Π_3 is also equal to zero by using Lemma 4. Therefore, we obtain from (65) that

$$\lim_{U \rightarrow \infty} \frac{\mathcal{C}_U}{N_t \log_2 \log_2 U} \geq (1 - P_\beta). \quad (69)$$

By applying Lemma 3, we obtain the lower bound of the asymptotic sum capacity in (31).

2) *Upper Bound for Asymptotic Sum Capacity:* We can bound the sum capacity for OSDMA-TF by that for the case of feedback from all users, denoted as \mathcal{C}^+ . Therefore,

$$\mathcal{C} \leq \mathcal{C}^+ = E \left[\max_{m=1, \dots, M} \sum_{n=1}^{N_t} \log_2 \left(1 + \max_{1 \leq u \leq U} \text{SINR}_u \right) \right]. \quad (70)$$

Note the difference in the subscript for the second “max” operator in the above equation from that of \mathcal{C} in (7). The case of feedback from all users is analyzed in [24]. Theorem 1 of [24] shows that

$$N_t \log_2 \log_2 U \leq \mathcal{C}^+ \leq N_t \log_2 \log_2 (UM). \quad (71)$$

Therefore,

$$\lim_{U \rightarrow \infty} \frac{\mathcal{C}^+}{N_t \log_2 \log_2 U} = 1. \quad (72)$$

From (70) and (72), we obtain the upper bound of the asymptotic sum capacity in (31). Thereby, we completes the proof of Theorem 2.

REFERENCES

- [1] “IEEE 802.16e amendment: Physical and medium access control layers for combined fixed and mobile operation in licensed bands,” *IEEE Standard 802.16*, 2005.
- [2] M. Costa, “Writing on dirty paper,” *IEEE Trans. on Info. Theory*, vol. 29, no. 3, pp. 439 – 441, 1983.
- [3] G. Caire and S. Shamai, “On the achievable throughput of a multiantenna Gaussian broadcast channel,” *IEEE Trans. on Info. Theory*, vol. 49, no. 7, pp. 1691–1706, 2003.
- [4] L.-U. Choi and R. Murch, “A transmit preprocessing technique for multiuser MIMO systems using a decomposition approach,” *IEEE Trans. on Wireless Communications*, vol. 3, no. 1, pp. 20–24, 2004.
- [5] K.-K. Wong, R. Murch, and K. Letaief, “A joint-channel diagonalization for multiuser MIMO antenna systems,” *IEEE Trans. on Wireless Communications*, vol. 2, no. 4, pp. 773–786, 2003.
- [6] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, “Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels,” *IEEE Trans. on Signal Processing*, vol. 52, no. 2, pp. 461 – 471, 2004.
- [7] G. Dimic and N. Sidropoulos, “Low-complexity downlink beamforming for maximum sum capacity,” in *Proc., Int. Conf. Acoust., Speech and Sig. Proc.*, vol. 4, 2004.
- [8] M. Schubert and H. Boche, “Solution of the multiuser downlink beamforming problem with individual SINR constraints,” *IEEE Trans. on Veh. Technol.*, vol. 53, pp. 18–28, Jan. 2004.
- [9] S. Serbetli and A. Yener, “Transceiver optimization for multiuser MIMO systems,” *IEEE Trans. on Signal Processing*, vol. 52, no. 1, pp. 214–226, 2004.

- [10] Y.-S. Choi, S. Alamouti, and V. Tarokh, "Complementary beamforming: new approaches," *IEEE Trans. on Communications*, vol. 54, no. 1, pp. 41–50, 2006.
- [11] R. Knopp and P. Humblet, "Information capacity and power control in single-cell multiuser communications," in *Proc., IEEE Intl. Conf. on Communications*, vol. 1, (Seattle, WA), pp. 331–335, 1995.
- [12] C. Swannack, E. Uysal-Biyikoglu, and G. W. Wornell, "MIMO broadcast scheduling with limited channel state information," in *Proc., Allerton Conf. on Comm., Control, and Computing*, Sept. 2005.
- [13] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE Journal on Sel. Areas in Communications*, vol. 24, no. 3, pp. 528–541, 2006.
- [14] T. Yoo, N. Jindal, and A. Goldsmith, "Finite-rate feedback MIMO broadcast channels with a large number of users." To appear Proc. of IEEE Intl. Symposium on Info. Theory, 2006.
- [15] W. Choi, A. Forenza, J. G. Andrews, and R. W. Heath Jr., "Opportunistic space division multiple access with beam selection," to appear *IEEE Trans. on Communications*.
- [16] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channels with partial side information," *IEEE Trans. on Info. Theory*, vol. 51, pp. 506–522, Feb. 2005.
- [17] Z. Shen, R. Chen, J. G. Andrews, R. W. Heath Jr., and B. L. Evans, "Low complexity user selection algorithms for multiuser MIMO systems with block diagonalization," to appear *IEEE Trans. on Signal Processing*, 2006.
- [18] D. Gesbert and M.-S. Alouini, "How much feedback is multi-user diversity really worth?," in *Proc., IEEE Intl. Conf. on Communications*, vol. 1, pp. 234–238, 2004.
- [19] V. Hassel, M.-S. Alouini, D. Gesbert, and G. Oien, "Exploiting multiuser diversity using multiple feedback thresholds," in *Proc., IEEE Veh. Technology Conf.*, vol. 2, pp. 1302–1306, 2005.
- [20] S. Sanaye and A. Nosratinia, "Exploiting multiuser diversity with only 1-bit feedback," in *Proc., IEEE Wireless Communications and Networking Conf.*, vol. 2, pp. 978–983, 2005.
- [21] S. Sanaye and A. Nosratinia, "Opportunistic beamforming with limited feedback," in *Proc., IEEE Asilomar*, pp. 648–652, Nov. 2005.
- [22] T. Tang and R. W. Heath, Jr., "Opportunistic feedback for downlink multiuser diversity," *IEEE Commun. Lett.*, vol. 9, no. 10, pp. 948–950, 2005.
- [23] T. Tang, R. W. Heath Jr., S. Cho, and S. Yun, "Opportunistic feedback in multiuser MIMO systems with linear receivers," submitted to *IEEE Trans. on Communications*, Nov. 2005.
- [24] K.-B. Huang, R. W. Heath Jr., and J. G. Andrews, "Joint beamforming and scheduling for SDMA systems with limited feedback," submitted to *IEEE Trans. on Communications*; available at ArXiv: <http://arxiv.org/abs/cs.IT/0606121>, June 2006.
- [25] D. J. Love, R. W. Heath Jr., W. Santipach, and M. L. Honig, "What is the value of limited feedback for MIMO channels?," *IEEE Comm. Mag.*, vol. 42, pp. 54–59, Oct. 2004.
- [26] D. J. Love and R. W. Heath Jr., "Feedback techniques for MIMO channels," in *MIMO Antenna Technology for Wireless Communications*, (Boca Raton, FL), CRC Press Inc, 2006.
- [27] L. Yang, M.-S. Alouini, and D. Gesbert, "Further results on selective multiuser diversity," in *7th ACM/IEEE Intl. Symp.*

on Model., Analysis and Sim. of Wireless and Mobile Sys., pp. 25–30, Oct. 2004.

[28] N. Jindal, “MIMO broadcast channels with finite rate feedback,” *submitted to IEEE Trans. Information Theory*, 2006.

[29] K. Zyczkowski and M. Kus, “Random unitary matrices,” *J. Phys.*, vol. A27, pp. 4235–45, 1994.

[30] D. J. Love, R. W. Heath Jr., and T. Strohmer, “Grassmannian beamforming for multiple-input multiple-output wireless systems,” *IEEE Trans. on Info. Theory*, vol. 49, pp. 2735–47, Oct. 2003.

[31] P. Ding, D. J. Love, and M. D. Zoltowski, “On the sum rate of channel subspace feedback for multi-antenna broadcast channels,” in *Proc., IEEE Globecom*, vol. 5, pp. 2699–2703, Nov. 2005.

[32] H. Alzer, “On some inequalities for the incomplete Gamma function,” *MATHEMATICS OF COMPUTATION*, vol. 66, no. 218, pp. 771–778, 2005.

[33] D. Bertsekas and R. Gallager, *Data networks*. Prentice Hall, 1992.

[34] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Prentice Hall, 2001.

[35] S. Janson, T. Luczak, and A. Rucinski, *Random Graphs*. John Wiley, 2000.